Gas Vesicle Genes Identified in *Bacillus megaterium* and Functional Expression in *Escherichia coli*

NING LI AND MAURA C. CANNON*

Department of Microbiology, University of Massachusetts, Amherst, Massachusetts 01003

Received 4 December 1997/Accepted 4 March 1998

Gas vesicles are intracellular, protein-coated, and hollow organelles found in cyanobacteria and halophilic archaea. They are permeable to ambient gases by diffusion and provide buoyancy, enabling cells to move upwards in liquid to access oxygen and/or light. In halobacteria, gas vesicle production is encoded in a 9-kb cluster of 14 genes (4 of known function). In cyanobacteria, the number of genes involved has not been determined. We now report the cloning and sequence analysis of an 8,142-bp cluster of 15 putative gas vesicle genes (*gvp*) from *Bacillus megaterium* VT1660 and their functional expression in *Escherichia coli*. Evidence includes homologies by sequence analysis to known gas vesicle genes, the buoyancy phenotype of *E. coli* strains that carry this *gvp* gene cluster, the presence of pressure-sensitive, refractile bodies in phase-contrast microscopy, structural details in phase-contrast microscopy, structural details in direct interference-contrast microscopy, and shape and size revealed by transmission electron microscopy. In *B. megaterium*, the *gvp* region carries a cluster of 15 putative genes arranged in one orientation; they are open reading frame 1 and *gvpA, -P, -Q, -B, -R, -N, -J, -F, -G, -L, -S, -K, -J, -T, and -U*, of which the last 11 genes, in a 5.7-kb gene cluster, are the maximum required for gas vesicle synthesis and function in *E. coli*. To our knowledge, this is the first example of a functional gas vesicle gene cluster in nonaquatic bacteria and the first example of the interspecies transfer of genes resulting in the synthesis of a functional organelle.

Gas vesicles are intracellular, hollow organelles found in many bacteria from aqueous environments and are most studied in cyanobacteria (8) and halophilic archaea (10, 13, 20). Gas vesicles are permeable to ambient gases by diffusion. The physiological role demonstrated for gas vesicles is that they provide buoyancy, enabling cells to move upwards in liquid to access oxygen and/or light (41). Gas vesicles range from 30 to 250 nm in width and from 50 nm to 1 μm in length (9, 41), but have a relatively constant size within each species. The morphology and main structural protein of gas vesicles is conserved among species. Gas vesicles are found both individually in cells and clustered together to form gas vacuoles that can be seen as refractile bodies in phase-contrast microscopy.

In the halophilic archaea *Halofex mediterranei* and * Halo bacterium salinarium* PHH1, a 9-kb cluster (*vac*) of 14 genes (*gvp*) is all that is required for gas vesicle production and regulation (15, 21). *H. mediterranei* has one chromosomal *vac* region (14, 15), while *H. salinarium* PHH1 has a chromosomal and a plasmid *vac* cluster of genes (15, 21, 31). The chromosomal *vac* is expressed only in the absence of the plasmid-borne *vac* region (22). *Halobacterium halobium* also has a chromosomal- and a plasmid-borne *vac* cluster of genes (33, 43). These plasmid and chromosome-borne *gvp* genes are in the same order as each other and are arranged divergently in two groups, *gvpDEFGHJKLM*, followed upstream by *gvpACNO* (11, 18, 21, 23). Another halophilic archaean, *Natronobacterium vacuolatum*, that forms gas vesicles has a different *vac* arrangement; seven of its *gvp* genes are clustered as *gvpACNOFGH*, that forms gas vesicles has a different *vac* arrangement; seven of its *gvp* genes are clustered as *gvpACNOFGH*, and are cotranscribed (27, 32). In cyanobacteria, the total number of genes involved in gas vesicle formation has not been determined. There are two copies of *gvpA* and one copy of *gvpC* identified in *Calothrix* sp. strain 7601 (6, 8).

In *Anabaena flos-aquae*, there are at least five copies of *gvpA* and one copy each of *gvpC, -N, -J, -K, -F*, and *-L* (19, 24).

The main structural protein of gas vesicles is GvpA. The 70-amino-acid, extremely hydrophobic protein was first cloned from *Calothrix* sp. strain PCC7601 (with an oligonucleotide based on the amino acid sequence) (37), and a probe based on this *gvpA* gene was subsequently used to clone *gvpA* from other cyanobacteria (7, 42) and halobacteria (13, 20, 36). All of the other cloned *gvp* genes were identified by being either homologous to or contiguous with *gvpA*. The highly conserved GvpA protein forms a linear crystalline array of ribs that make up the cylindrical shell and conical ends of the gas vesicle, while GvpC is located on the outer surfaces and adds strength and shape (3, 18, 30, 41). In *H. halobium* and *H. salinarium*, a nonessential homolog of GvpA named GvpB is located outside the well-studied *gvp* cluster (12, 20). GvpD and GvpE were shown to have a regulatory role in *gvp* gene expression (25, 29, 34). The functions of the other 10 gas vesicle gene products are still unknown. Apart from GvpA, GvpB, and, to a lesser extent, GvpN, sequence conservation of *gvp* gene products between genera is low.

*Bacillus megaterium* is generally considered to be a soil bacterium, although it has been found in diverse environments (38). Since gas vesicles have been described exclusively in bacteria from aqueous habitats and have not been found in the bacilli, the discovery of a functional gas vesicle gene cluster in *B. megaterium* is novel. Open reading frames (ORFs) with homologies to *gvp* genes were identified in *B. megaterium* VT1660 in the course of a screening of Tn917-LTV1 (4) transposon banks for polyhydroxyalkanoic acid mutants. One polyhydroxyalkanoic acid-overproducing mutant had the transposon inserted at a site contiguous with the *gvp* genes. In this paper, we report the discovery of a cluster of *gvp* genes in *B. megaterium*. We describe the cloning and sequence analysis of an 8,142-bp DNA fragment encoding a cluster of *gvp* genes that when transferred to *Escherichia coli*, conferred a buoyancy

* Corresponding author. Mailing address: Department of Microbiology, University of Massachusetts, Amherst, MA 01003. Phone: (413) 545-0092. Fax: (413) 545-1578. E-mail: mcannon@bio.umass.edu.
TABLE 1. Strains and plasmids used in this study

<table>
<thead>
<tr>
<th>Strain or plasmid</th>
<th>Relevant characteristics*</th>
<th>Source or reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Strains</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>E. coli</em> DH5α</td>
<td>deoR endA1 gvrA69 hsdR17 (F− mK−) recA1 relA1 supE44 thi-1 Δ(lacZΔX74-argF59)</td>
<td>Clontech</td>
</tr>
<tr>
<td><em>B. megaterium</em></td>
<td>Part of the culture collection of the University of Texas, J. W. Foster wild-type strain</td>
<td>ATCC*</td>
</tr>
<tr>
<td>VT1660</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>B. megaterium</em></td>
<td>Arginine auxotroph, plasmidless derivative of <em>B. megaterium</em> VT1660, Em' Lm'</td>
<td>This study</td>
</tr>
<tr>
<td>B001S</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11561</td>
<td>Wild type</td>
<td>ATCC</td>
</tr>
<tr>
<td><strong>Plasmids</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pBluescriptIISK</td>
<td>Cloning vector, Amp′</td>
<td>Stratagene</td>
</tr>
<tr>
<td>pBluescriptIISK</td>
<td>Cloning vector, Amp′</td>
<td>Stratagene</td>
</tr>
<tr>
<td>pHPS9</td>
<td>Bacillus-*E. coli shuttle vector, Em′ Cm′</td>
<td>17</td>
</tr>
<tr>
<td>pNL4</td>
<td>IR-L end of Tn917-LTV1 and 6.25-kb contiguous <em>B. megaterium</em> B001S sequences</td>
<td>This study</td>
</tr>
<tr>
<td>pNL21</td>
<td>6.4-kb SacI-BamHI fragment of pNL4, cloned in SacI-BamHI of pBluescriptIISK</td>
<td>This study</td>
</tr>
<tr>
<td>pNL22</td>
<td>2.6-kb PstI fragment of <em>B. megaterium</em> VT1660 genomic DNA, cloned in PstI site of pBluescriptIISK</td>
<td>This study</td>
</tr>
<tr>
<td>pNL24</td>
<td>8,522-bp composite of pNL21 and pNL22, cloned in pBluescriptIISK, all putative gvp genes present</td>
<td>This study</td>
</tr>
<tr>
<td>pNL25</td>
<td>EcoRI in ORF1 to PstI in araC, cloned into EcoRI-PstI of pBluescriptIIKS</td>
<td>This study</td>
</tr>
<tr>
<td>pNL25-SK</td>
<td>Same as pNL25, except cloned in pBluescriptIISK</td>
<td>This study</td>
</tr>
<tr>
<td>pNL26</td>
<td>HindIII in ORF1-yrpA intergenic region to PstI in araC, cloned into HindIII-PstI of pBluescriptIISK</td>
<td>This study</td>
</tr>
<tr>
<td>pNL26-SK</td>
<td>Same as pNL26, except cloned in pBluescriptIISK</td>
<td>This study</td>
</tr>
<tr>
<td>pNL27</td>
<td>EcoRI in gvpA to PstI in araC, cloned into EcoRI-PstI of pBluescriptIISK</td>
<td>This study</td>
</tr>
<tr>
<td>pNL28</td>
<td>EcoRV in gvpP to PstI in araC, cloned into EcoRV-PstI of pBluescriptIISK</td>
<td>This study</td>
</tr>
<tr>
<td>pNL29</td>
<td>HindIII in gvpQ to PstI in araC, cloned into HindIII-PstI of pBluescriptIISK</td>
<td>This study</td>
</tr>
<tr>
<td>pNL30</td>
<td>BglII in gvp-yrpB intergenic region to PstI in araC, cloned into BglII-PstI of pBluescriptIISK</td>
<td>This study</td>
</tr>
<tr>
<td>pNL32</td>
<td>HindIII in ORF1-yrpA intergenic region to HindIII in gvpU, cloned into HindIII of pBluescriptIISK, orientation as in pNL24</td>
<td>This study</td>
</tr>
<tr>
<td>pNL20</td>
<td>6.4-kb SacI-BamHI fragment of pNL4, cloned in SacI-BamHI of pHPS9</td>
<td>This study</td>
</tr>
<tr>
<td>pNL40</td>
<td>8.5-kb SacI-BamHI in vector, cloned in SacI-BamHI of pHPS9, all putative gvp genes present</td>
<td>This study</td>
</tr>
</tbody>
</table>

* Em′, erythromycin resistant; Lm′, lincomycin resistant; Cm′, chloramphenicol resistant; Amp′, ampicillin resistant.

* ATCC, American Type Culture Collection.

phenotype on its host resulting from the synthesis of gas vesicles.

**MATERIALS AND METHODS**

**Bacterial strains and plasmids.** The bacterial strains and plasmids used in this study are listed in Table 1.

**Media and growth conditions.** Cultures were grown at 37°C in liquid media, aeration by rotation at 200 rpm in either Luria-Bertani (LB) broth (28) or M9 minimal salts (Life Technologies) with 1% (w/v) glucose. Cultures were also grown on LB broth and M9 minimal salts-1% glucose plates containing 1.2% agar (A4550 [Sigma]) at 37°C. To increase sporulation, *B. megaterium* was grown on sporulation medium, which contained 0.8% nutrient broth (Difco), 0.3% yeast extract (Difco), and 25 μg of MnSO₄ · H₂O per ml. For growth of *B. megaterium* VT1660 and derivatives, minimal medium was supplemented with arginine HCl at 50 μg/ml (wt/vol). For plasmid selections, the appropriate antibiotics were included in the media: ampicillin (100 μg/ml [AMP][3]) or erythromycin (400 μg/ml [EM][3]) for plasmid selection in *E. coli* and chloramphenicol (12 μg/ml [CM][2]) or erythromycin (1 μg/ml [EM][1]) and lincomycin (25 μg/ml [LM][2]) for plasmid selection in *Bacillus* strains.

**Transformations.** *E. coli* was transformed by electroporation of competent cells with the *E. coli* pulser (Bio-Rad Laboratories) and according to the manufacturer’s instructions. *B. megaterium* was transformed by a polyethylene glycol-mediated transformation method (16) with protoplasting and hypertonic media as previously described (5, 39).

**Cloning of the gvp region.** Purification of genomic and plasmid DNA, Southern blotting, colony hybridization, and dephosphorylation of DNA vectors were performed by standard procedures (35). To clone DNA sequences contiguous with the transposon in the transposant, *B. megaterium* B001S genomic DNA was cut with BamHI self-ligated, and transformed into *E. coli*. Following selection on LB-AMP[106] plates, colonies were screened. The plasmid pNL4 carried the left inverted repeat (IR-L) end of Tn917-LTV1 and flanking chromosomal DNA (Fig. 1). The SacI-BamHI fragment of pNL4 was ligated into the SacI-BamHI sites of plasmid pBluescriptIISK. Following transformation of *E. coli* and selection on LB-AMP[106], transformants were screened, thus yielding the plasmid pNL21. Chromosomal DNA sequences, overlapping and contiguous with pNL21 and distal to IR-L, were cloned from *B. megaterium* B001S genomic DNA. The PstI fragment size was identified in a Southern blot by using a 32P-labeled synthetic oligonucleotide probe (5′-TCGGTTGAAACGTTGTGTC-3′) homologous to gvpS. The approximate-size fragments were excised from an agarose gel, extracted with GeneClean (Bio 101) and ligated into PstI-cut, dephosphorylated pBluescriptIISK. White colonies on LB-AMP[106] plates, supplemented with 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal) and isopropyl-β-D-thiogalactopyranoside (IPTG), were screened by colony hybridization and a 5′-end-labeled probe homologous to gvpS. A plasmid with native DNA sequences of gvp genes was constructed as follows. pNL22 was completely cut with NotI and partially cut with BamHI. 2.3-kb NotI-BamHI fragments were then excised from an agarose gel, extracted with GeneClean, and ligated into the NotI-BamHI-linearized plasmid pNL21. The plasmid thus constructed, pNL24, was confirmed for accuracy across the BamHI junction by sequencing with the oligonucleotide probe (gvpS) described above used as a primer. Plasmids pNL25, pNL26, pNL27, pNL28, pNL29, and pNL30 are subclones of pNL24 in pBluescriptIISK, unless otherwise noted (Table 1).

Cloned genomic DNA was subcloned into pHPS9 for expression in *B. megaterium*. The 6.4-kb SacI-BamHI fragment of pNL4 was subcloned from pNL4 into the SacI-BamHI sites of pHPS9, generating pNL30. This plasmid carries part of the gvp cluster. The complete cluster of cloned gas vesicle genes were cloned in pHPS9 by the digestion of pNL24 completely with SacI and partially with BamHI, followed by extraction of 8.5-kb SacI-BamHI fragments from agarose with GeneClean and ligation into the SacI-BamHI sites of pHPS9. The plasmid construct was confirmed by restriction mapping and was named pNL40.
Sequencing of the *gvp* region.

DNA fragments were subcloned into pBlue-ScriptIISK, and sets of nested deletion clones from both sides were generated with the Erase-a-Base system (Promega). DNA was sequenced with USB Sequenase, version 2 (Amersham Life Science). Sequence assembly and analysis were performed with Lasergene (DNAStar, Inc.). Homologies to known genes were determined by Gapped BLAST and PSI BLAST sequence analysis programs (1) of the National Center for Biotechnology Information at the National Library of Medicine.

Buoyancy test.

Cells were grown on 100-mm-diameter plates in LB broth with ampicillin and IPTG for 24 h at 37°C. The cells from each plate were resuspended in 10 ml of saline in 13-mm-diameter test tubes. The tubes were left stationary and undisturbed at room temperature for at least 16 h, at which time the cell buoyancy was determined by the visual degree of turbidity of the culture medium.

Phase-contrast microscopy. Wet mounts of cultures were visualized at ×1,000 magnification with a light microscope with phase-contrast attachments (Labophot-2; Nikon, Inc.).

DIC microscopy. Wet mounts of cultures were visualized at ×1,000 magnification in a light microscope with Nomarski attachments for direct interference contrast (DIC) microscopy (Labophot-2).

Electron microscopy. E. coli cells grown on LB broth with ampicillin and IPTG were resuspended in 10 mM Tris HCl (pH 7) and incubated at 25°C for 30 min with lysozyme (2 mg per ml) to generate protoplasts. To lyse these protoplasts, sodium dodecyl sulfate was added to a final concentration of 0.2%. Samples taken from the top of the supernatants of cell lysates and samples of protoplasts of E. coli were analyzed for gas vesicles with a Philips transmission electron microscope (CM10) at an 80-kV acceleration voltage. Drops of samples were adsorbed onto Formvar- and carbon-coated 400-mesh copper grids, negatively stained with 1% (wt/vol) uranyl acetate in water for 30 to 60 s, and blotted dry.

Pressure sensitivity test of refractile bodies. Samples of overnight cultures grown in LB broth with ampicillin and IPTG were centrifuged at 16,000 × g for 5 to 40 min in a microcentrifuge. Pellets were gently resuspended in saline, and cells were visualized by phase-contrast microscopy for the presence of refractile bodies or left at room temperature for 16 h to test for effects on cell buoyancy.

**RESULTS**

*B. megaterium* *gvp* gene cluster analysis. *E. coli* synthesized functional gas vesicles when carrying an 8,142-bp region of DNA cloned from *B. megaterium* B001S (data below). Nucleotide and predicted amino acid sequence analyses of this region revealed a cluster of 1 partial and 15 complete, putative genes, as depicted on plasmid pNL24 in Fig. 1A.

**FIG. 1.** The *gvp* gene cluster and contiguous sequences in *B. megaterium* B001S. (A) [from top to bottom] Tn917-LTV1 and contiguous sequences in the chromosome of strain B001S and relevant restriction enzyme sites; cloned fragments of DNA in pNL4, pNL21, and pNL22 contiguous with the IR-L sequences; cloned DNA fragment in pNL24, map of putative genes, intergenic regions in base pairs (igrs), and relevant restriction enzyme sites; and ruler of sequence in base pairs. (B) Subclones of pNL24 and their ability to produce gas vesicles in *E. coli*. Gas vesicles were observed as refractile bodies by phase-contrast microscopy. ++++++, gas vesicles throughout all cells in culture; +++, +++, slightly reduced quantity of gas vesicles in most cells; +++, gas vesicles in approximately 50% of cells and in reduced quantity; +, very few gas vesicles in about 10% of cells; -, no gas vesicles observed. IR-L and IR-R, inverted repeats at lacZ end (left) and Tn917 end (right), respectively; ColE1, *E. coli* origin of replication; lacZ, β-galactosidase gene; cat, chloramphenicol resistance gene; bla, ampicillin resistance gene; erm, erythromycin resistance gene.
of a map of this gvp region was based on a sequence analysis that revealed ORFs, putative ribosome binding sites, and homologies to known and putative genes in data banks (Table 2). These 15 putative genes read in one direction. Nine of the 15 have amino acid sequence similarity to known and putative gvp gene products, as determined by BLAST searches, while the remaining six have no significant homology to known genes. To comply with standard nomenclature, a four-letter designation was assigned to each putative gene according to the greatest homologies. Where paralogs to known genes exist, the gene of lesser homology to known genes was assigned a new letter, as were the six putative genes with no homology to known genes.

GvpA and GvpB have high homology to each other and to other known GvpA proteins (Fig. 2A). GvpS and GvpJ have

### Table 2. Sequence analysis results

<table>
<thead>
<tr>
<th>Gene name</th>
<th>No. of amino acids</th>
<th>Mol mass (Da)</th>
<th>Isoelectric point</th>
<th>Homologies to known and putative genes (accession no.)</th>
<th>BLAST score</th>
</tr>
</thead>
<tbody>
<tr>
<td>ORF1</td>
<td>161</td>
<td>16,147</td>
<td>5.39</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td>gvpA</td>
<td>86</td>
<td>9,291</td>
<td>4.37</td>
<td>GVPA_MICBC (P08412), GVPA_APHFL (P10397)</td>
<td>95.3, 95.5</td>
</tr>
<tr>
<td>gvpP</td>
<td>161</td>
<td>17,945</td>
<td>4.89</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td>gvpQ</td>
<td>157</td>
<td>17,583</td>
<td>9.40</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td>gvpB</td>
<td>88</td>
<td>9,618</td>
<td>4.44</td>
<td>GVPA_MICBC (P08412), GVPA_APHFL (P10397)</td>
<td>94.6, 92.8</td>
</tr>
<tr>
<td>gvpR</td>
<td>88</td>
<td>10,344</td>
<td>5.18</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td>gvpN</td>
<td>308</td>
<td>34,700</td>
<td>5.19</td>
<td>GVPN_APHFL (P55150), GVPN_HALME (P02239), GVPN_HALHA (P33952), GVPN_HALSA (P33965)</td>
<td>228.0, 178.0, 178.0, 178.0</td>
</tr>
<tr>
<td>gvpF</td>
<td>255</td>
<td>28,864</td>
<td>5.15</td>
<td>GVPF_APHFL (P55149), GVPF_HALME (P02231), GVPF_HALHA (P24370)</td>
<td>67.0, 65.5, 63.0</td>
</tr>
<tr>
<td>gvpG</td>
<td>88</td>
<td>10,540</td>
<td>4.60</td>
<td>GVPG_HALHA (P24371), GVPG_HALSA (P33960)</td>
<td>33.9, 32.8</td>
</tr>
<tr>
<td>gvpL</td>
<td>269</td>
<td>31,253</td>
<td>5.02</td>
<td>GVPL_HALHA (P24376), GVPL_HALME (P02237), GVPL_HALSA (P24370)</td>
<td>68.5, 61.5, 50.1, 48.2</td>
</tr>
<tr>
<td>gvpS</td>
<td>95</td>
<td>10,422</td>
<td>4.53</td>
<td>GVPS_APHFL (P55147), GVPS_PSEAN (P22453), GVPS_APHFL (P10397)</td>
<td>37.6, 31.7, 31.3</td>
</tr>
<tr>
<td>gvpK</td>
<td>94</td>
<td>10,448</td>
<td>4.18</td>
<td>GVPK_APHFL (P55148), GVPK_HALME (P24375), GVPK_HALHA (P24370), GVPK_HALSA (P33963)</td>
<td>65.2, 58.9, 55.6, 53.8</td>
</tr>
<tr>
<td>gvpJ</td>
<td>100</td>
<td>11,261</td>
<td>4.95</td>
<td>GVPJ_APHFL (P55147), GVPJ_APHFL (P10397), GVPJ_HALME (P02235), GVPJ_HALHA (P07060), GVPJ_HALSA (P33956)</td>
<td>107.0, 49.7, 49.7, 49.4, 48.6, 48.6, 48.2, 47.9, 39.4</td>
</tr>
<tr>
<td>gvpT</td>
<td>292</td>
<td>32,739</td>
<td>4.26</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td>gvpU</td>
<td>138</td>
<td>14,960</td>
<td>5.72</td>
<td>None</td>
<td></td>
</tr>
<tr>
<td>araC</td>
<td>49</td>
<td>5,831</td>
<td>Not done</td>
<td>HTH_ARAC (AB001488)</td>
<td>60.4</td>
</tr>
</tbody>
</table>

* Loci and accession numbers are SWISS-PROT unless otherwise stated. None, no discernible similarity to known sequences.

b Partial protein.

c DDBJ.
26% identity to each other and less sequence similarity to other known GvpJ proteins. *B. megaterium* GvpA and GvpB show sequence similarity to GvpS and GvpJ and less homology to the low-molecular-weight protein GvpK (Fig. 2B). As well as sequence similarity, GvpA, -B, -S, and -J have low molecular weights and hydrophathy profiles that show the presence of a center hydrophobic domain, bordered at the C and N termini by hydrophilic amino acid sequences. GvpF and GvpL have 24% identity to each other and both show low sequence similarities to GvpF and GvpL from both the cyanobacteria and halophilic archaea. GvpN and GvpG have no homology to other known GvpJ proteins. While this pattern of amino acid identity is insufficient for demonstration of sequence similarity, GvpA, GvpB, GvpJ, GvpK, and GvpL carry genes whose predicted amino acid sequences show homology to known *gvp* genes was tested for gas vesicle function in *E. coli*. This experiment was carried out by first cloning the 8,142-bp region of the *B. megaterium* chromosomal DNA that carries genes whose predicted amino acid sequences show homology to known *gvp* genes, deletion of *gvp* genes was tested for gas vesicle function in *E. coli* (pNL24) cells re-suspended in the medium rather than sinking to the bottom of the tube, as was the case with the control *E. coli* (pBlue-scriptIIKS) cells. Following overnight growth at 37°C in LB broth with ampicillin and IPTG, *E. coli* (pNL24) and *E. coli* (pBlue-scriptIIKS) were tested for a buoyancy phenotype and for the presence of gas vesicles as was *E. coli* (pBlue-scriptIIKS) cells. Following overnight growth at 37°C, cells carrying pNL24 also showed the presence of refractile bodies in phase-contrast microscopy and hollow-looking structures in DIC microscopy, while no such structures were present in the control. These results were consistent with the 8,124-bp region of *B. megaterium* DNA coding for the synthesis of functional gas vesicles. In order to determine the minimum length of DNA required for the putative gas vesicle synthesis, deletion derivatives of pNL24 were constructed, and *E. coli* strains carrying these plasmids were tested for buoyancy phenotype and for the presence of gas vesicles as was *E. coli* (pNL24). The plasmid constructs are described in Table 1. The results of phase-contrast and DIC microscopy gave an approximation of the quantities of presumptive gas vesicles in the cells and are summarized in Fig. 1B. The results of both types of microscopy were in agreement with each other and with the results of the

FIG. 2. Homologies of predicted amino acid sequences. (A) Multiple sequence alignment of *B. megaterium* GvpA and GvpB with representative homologs from cyanobacteria and the halophilic archaea, with the majority consensus sequences shown in black. (B) Multiple sequence alignment of *B. megaterium* GvpA, GvpB, GvpJ, GvpK, with multiple consensus sequences shown in black. The Clustal method with the PAM250 residue weight table was used. GVPA_APHFL, GvpA of *H. halobium*; GvpA1_HALHA, GvpA (plasmid borne) of *H. halobium*; GVP_A1_HALHA, GvpA of *H. halobium*; Bmeg, *B. megaterium*. (C) Pairwise alignment of *H. halobium* GvpQ with *B. megaterium* GvpQ, with amino acid identities shown in black. The Clustal method with the PAM250 residue weight table was used. GVPA_APHFL, GvpA of *H. halobium*. 

There are very few or no intergenic sequences between 14 of the putative gene products, from GvpA to GvpU, while GvpS and -K overlap by 47 bp (Fig. 1A). Upstream from GvpA is a 442-bp region with no discernible coding function. This region has a G + C ratio of 25%, compared to 34 to 45% for the putative *gvp* coding sequences, and possibly carries a promoter for *gvpA* and downstream genes. Upstream of this intergenic region is ORF1, which is preceded by 63 bp on this cloned *gvp* region of *B. megaterium* DNA. A study of the regulation of *gvp* gene expression is in progress, but culture conditions that result in a buoyancy phenotype of *B. megaterium* have not been determined. 

*B. megaterium* gvp gene cluster functions in *E. coli*. The 8,142-bp region of the *B. megaterium* chromosomal DNA that is contiguous with Tn917-LTV1 in the mutant B001S and that carries genes whose predicted amino acid sequences show homology to known *gvp* genes was tested for gas vesicle function in *E. coli*. This experiment was carried out by first cloning the 8,522-bp SalI-PstI fragment from B001S in pBluescriptIIKS to make pNL24 as described above. Following overnight growth at 37°C on LB broth with ampicillin and IPTG, *E. coli* (pNL24) and *E. coli* (pBlue-scriptIIKS) were tested for a buoyancy phenotype and for the presence of gas vesicles as was *E. coli* (pBlue-scriptIIKS) cells. Following overnight growth at 37°C, cells carrying pNL24 also showed the presence of refractile bodies in phase-contrast microscopy and hollow-looking structures in DIC microscopy, while no such structures were present in the control. These results were consistent with the 8,124-bp region of *B. megaterium* DNA coding for the synthesis of functional gas vesicles. In order to determine the minimum length of DNA required for the putative gas vesicle synthesis, deletion derivatives of pNL24 were constructed, and *E. coli* strains carrying these plasmids were tested for buoyancy phenotype and for the presence of gas vesicles as was *E. coli* (pNL24). The plasmid constructs are described in Table 1. The results of phase-contrast and DIC microscopy gave an approximation of the quantities of presumptive gas vesicles in the cells and are summarized in Fig. 1B. The results of both types of microscopy were in agreement with each other and with the results of the
buoyancy test. The turbidity (buoyancy) of the cultures on standing at room temperature for 16 h and longer, as observed by the naked eye, was directly related to the quantities of presumptive gas vesicles in the cells at time zero (following overnight growth), as determined by phase-contrast and DIC microscopy. The buoyancy phenotypes are shown in Fig. 3, and the DIC microscopy results are demonstrated in Fig. 4. Figure 3 shows the results for E. coli(pNL25), E. coli(pNL26), and E. coli(pNL29) compared to E. coli(pBluescriptIIKS) after 16 h of standing at room temperature. The cultures carrying the B. megaterium cloned fragments remained dispersed throughout the medium, whereas the culture carrying the cloning vector alone sank to the bottom of the tube. This buoyancy phenotype was maintained for at least 3 days and was not observed thereafter. When the cultures with buoyancy phenotypes were disturbed by shaking after 16 h of standing at room temperature and then allowed to stand for a further 16 h, the buoyancy phenotype was reestablished and maintained for at least 3 days.

The shortest fragment of DNA shown to produce the presumptive gas vesicles in E. coli was the 6,040-bp HindIII-PstI fragment in pNL29. This fragment carries GvpB, -R, -N, -F, -G, -L, -S, -K, -J, -T, and -U and parts of GvpQ and AraC. We hypothesize, therefore, that these 11 complete genes are the maximum number of B. megaterium genes required by E. coli to produce functional gas vesicles. Since each of the genes within this 6,040-bp fragment has not been deleted, it is not yet possible to say that all of these 11 genes are necessary. GvpU and GvpB are required to produce gas vesicles under the conditions tested, based on the fact that pNL32 and pNL30, respectively, gave negative results for synthesis of gas vesicles in E. coli. Whether or not GvpA can substitute for GvpB has not been tested. In E. coli, the plasmid pNL29, with its shorter fragment of cloned DNA, produced the most gas vesicles and the most pronounced buoyancy phenotype of all of the plasmids tested.

The plasmids that gave rise to gas vesicles had their DNA fragment inserted in the cloning vector in the same orientation as the lacZ promoter of pBluescriptIIKS. In pNL25-SK and pNL26-SK, in which the inserted fragment of DNA reads in the opposite orientation to the lacZ promoter, no gas vesicles and no buoyancy phenotypes were observed, indicating that expression of at least some, if not all, of the gvp genes on pNL24 were transcribed from the lacZ promoter. Given these results, it is not surprising that pNL24 and pNL25 produced fewer gas vesicles than pNL26, due to a possible reduction in the rate of transcription at the end of ORF1. If this were the only factor determining the quantity of gas vesicles produced by these plasmids, one would not expect pNL29 to produce more gas vesicles than pNL26, pNL27, and pNL28. The results show otherwise (Fig. 1B), indicating that GvpF and/or GvpQ may play a regulatory role. A comparison of the results for pNL27 and pNL28 indicates that GvpP has no regulatory function under the experimental conditions used, while a comparison of the results for pNL28 and pNL29 shows clearly that the absence of GvpQ makes a significant difference in the quantity of gas vesicles synthesized, suggesting that GvpQ could be a negative regulator of gas vesicle synthesis. However, these data give no indication of the mode or mechanism by which such regulation may occur.

That the refractile bodies observed by phase-contrast microscopy are hollow vesicles is supported by the fact that the quantity of refractile bodies in E. coli(pNL26) decreased in proportion to centrifugation times (Fig. 5). Following 30 min of centrifugation at 16,000 × g, refractile bodies were present in approximately 10% of cells compared with those seen in 60% of cells prior to centrifugation. Loss of refractile bodies coincided with loss of buoyancy.

Electron microscopy identifies gas vesicles in E. coli. Transmission electron microscopy of E. coli(pNL26) showed structures within protoplasts that resembled the shape and size of gas vesicles found in cyanobacteria and halophilic archaea (Fig. 6). The gas vesicles in E. coli(pNL26) ranged from 25 to
70 nm in width and from 60 to 110 nm in length, but most of the gas vesicles observed were approximately 50 nm in width and 75 nm in length. A rib structure, as has been previously described for gas vesicles from *E. coli*, can be observed around the width of the gas vesicles from *E. coli*. Also, the gas vesicles appeared to be cylindrical with cone-shaped ends; those gas vesicles that appeared round were possibly being viewed from one end. Gas vesicles, both released and in cells, were found mainly in clusters, known as gas vacuoles, and are consistent with the structures that were observed as refractile bodies in phase-contrast microscopy and as hollow structures in DIC microscopy. In keeping with this is our observation by phase-contrast microscopy that the gas vesicles (as gas vacuoles) clustered together over a 15-min period following lysis of the protoplasts and release of gas vesicles. This observation is consistent with the known hydrophobicity of the gas vesicle structural protein, GvpA.

**Expression of a gvp cluster in *B. megaterium*.** Buoyancy phenotypes have not been observed in (i) *B. megaterium* B001S, the transposon mutant from which the gvp cluster was cloned; (ii) *B. megaterium* VT1660, the progenitor of the mutant B001S; (iii) *B. megaterium* 19213, the progenitor of VT1660; or (iv) *B. megaterium* 11561. These strains, with and without the plasmids pNL20 and pNL40, were cultured with a wide range of media and culture conditions, but gas vesicles could not be confirmed by the methods described above for *E. coli*, because they were observed in very small quantities and inconsistently. Since the plasmid pNL40 carries the 8,142-bp cluster of functional gvp genes, as demonstrated in *E. coli*, at a plasmid copy number of approximately five per cell, it is probable that the expression of gvp genes in *B. megaterium* is stringently regulated. The *B. megaterium* strains used in this study were motile. Motility was greatest in fresh cultures, while in late-stationary-phase cultures, the cells showed very little or no motility.

![Graph](http://jb.asm.org/)

**FIG. 5.** Pressure sensitivity of refractile bodies in *E. coli* (pNL26). See Materials and Methods for details. Triplicate samples of the same culture had a 5% margin of error.

![Micrographs](http://jb.asm.org/)

**FIG. 6.** Electron micrographs of *E. coli* (pNL26). (A) Protoplast of a dividing cell shows gas vesicles within the cell. Bar, 123 nm. (B) Gas vesicles from cell lysate. Bar, 44 nm.
gvp genes in other Bacillus strains. Sequences with high homology to gvpA and gvpS were identified in *B. megaterium* 11561 by Southern hybridization with gvpA and gvpS probes. This result indicated that at least one other strain of *B. megaterium* as well as strain 19213 carries gvp genes. A similar analysis of *B. subtilis* 168 showed no homology to the *B. megaterium* gvp gene probes. Genome sequence data for *B. subtilis* 168 have since confirmed the absence of gvp genes in this strain (26).

### DISCUSSION

The only known physiological role for gas vesicles is that they provide buoyancy to cells for vertical movement in liquid, and thus cells can position themselves at a depth that allows other metabolic activities to function. With gas vesicles, cells are known to position themselves in liquid for optimal light and oxygen. Gas vesicles have been previously described in microorganisms from aqueous environments, but have not been previously identified in *B. megaterium*. However, it is not surprising that gas vesicles should occur in soil organisms. It is normal for soil to become wet or flooded, and under such conditions, bacteria with gas vesicles and/or other means of motility could avoid being washed down deep into the earth, becoming depleted in the upper layers of soil. Since *B. megaterium* is an obligate aerobe and motile, the survival advantage of gas vesicles over other motility methods could be due to the lack of energy requirements by gas vesicles for function. Since older cultures of the *B. megaterium* strains used in this study had reduced motility, it is reasonable to speculate that older cells, especially large cells with gas vesicles, could have a survival advantage. Since a functional cluster of gvp genes is present in *B. megaterium*, as demonstrated by expression in *E. coli*, it is reasonable to assume that this gene cluster functions to provide buoyancy to *B. megaterium* under certain environmental conditions not yet defined. In size and shape, the *B. megaterium* gas vesicles synthesized in *E. coli* are similar to those described in cyanobacteria and halophilic archaea. Gas vesicle synthesis has not been reported in *E. coli* with known gvp genes.

It has been suggested that many of the gas vesicle proteins with possible structural function in the archaea may play a role in gas vesicle assembly, but the roles of these proteins have not been described in detail. GvpA, the main structural protein of gas vesicles, and its paralog, GvpB, have a highly conserved amino acid sequence, indicating that the physical structure of gas vesicles is conserved. GvpC, a minor structural protein involved in shaping and strengthening gas vesicles, has been identified in both cyanobacteria and archaea, but a homolog of GvpC has not been identified in *B. megaterium* gvp gene cluster. GvpB, -F, -G, -H, -I, -J, -K, -L, -M, and -N are possibly minor structural proteins of gas vesicles in halobacteria (10, 41). Of these genes, only GvpJ, -K, -F/L, and -N have been identified in *B. megaterium*. *B. megaterium* homologs of GvpB, -F, -G, -J, -K, -L, and -N are described in this paper, but homologs of GvpH, -I, and -M, which are possibly minor structural proteins, and GvpD, -E, and -O, which have possible regulatory functions, have not been identified. It is not surprising that none of the gvp gene products with regulatory functions has been identified, because regulation in *B. megaterium* may be very different from that in the archaea.

Nonorthologous functional equivalents of GvpH, -I, and -M may be among the *B. megaterium* GvpF, -Q, -R, -T, and -U proteins. The gvp cluster provides an interesting opportunity to examine the possibility of nonorthologous functional alternatives thus providing a functional identity to unclassified genes. In the *B. megaterium* gvp cluster, 38% of the ORFs (5 in 14, or 6 in 15 [including ORF1]) show no significant homology to known genes. The *E. coli* genome has 1,632 hypothetical and unclassified ORFs, based on sequence homology, which amounts to 38% of the total genome sequence (2). In *B. subtilis*, the function of 42% of the hypothetical genes cannot be predicted by homology to genes of known function (26). Nonorthologous gene alternatives could provide new insights not only into gene function but also into the evolutionary origins of organisms. In a comparison of the *B. megaterium* and *H. halobium* predicted gvp gene products, it is notable that GvpQ and GvpI have a number of features in common. GvpQ (17.6 kDa) has an isoelectric point of 9.4, while all of the other putative gvp gene products from *B. megaterium* have isoelectric points ranging from 4.2 to 5.7 (Table 2). The gvp gene products of halophilic archaea have isoelectric points ranging from 3.9 to 4.9, with the exception of GvpI (16 KDa) (21), which has an isoelectric point of 10.8. This fact, together with the similarity of their sizes, their sequence identity (at intervals), and similar hydropathy profiles, suggests that GvpQ of *B. megaterium* and GvpI of *H. halobium* could be nonorthologous functional equivalents.

We have identified a maximum of 11 genes in a 5.7-kb cluster from *B. megaterium* that are required for gas vesicle formation and function in *E. coli*. They are gvpB, -R, -N, -F, -G, -L, -S, -K, -J, -T, and -U. As well as these 11 genes, we have provided evidence that 3 and possibly 4 additional genes may be involved in gas vesicle formation; they are gvpA, -P, and -Q, and less likely, ORF1. The presence of paralogous genes in the gvp cluster of *B. megaterium*, which are also present in the gvp clusters in archaea and cyanobacteria, is interesting from both functional and evolutionary perspectives. This discovery of the *B. megaterium* gvp cluster and its functional expression in *E. coli* will enable the study of gas vesicle biogenesis, including assembly, gene product function, and regulation.

### ACKNOWLEDGMENTS

We thank Frank Cannon for critically reading the manuscript and Curt Thorne for the *Bacillus* strains. We also thank Tracy Guillemette and Heather Baker for excellent technical assistance and Lucy Yin for contributions to electron microscopy and photography.

This research was supported in part by a grant from the National Science Foundation (MCB 97-28066). The University of Massachusetts, Amherst, Central Microscopy Facility is supported by a grant from the NSF (NSF BBS 8714235).

### REFERENCES


Downloaded from http://jb.asm.org on October 14, 2017 by guest